Distribuzioni continue non uniformi

Simulazione & Logistica, I modulo
Lezione n. 9

Corso di Laurea in Informatica Applicata
Università di Pisa, sede di La Spezia
A.a. 2008/09, I semestre

Contenuti

- Variabili casuali, distribuzioni discrete, generatori
- Distribuzioni continue
- Generazione di distribuzioni non uniformi
- Generazione di distribuzioni esponenziali, normali e Weibull
- In pratica: generatori in GS DSLibs

Variabili casuali

- Spazio di probabilità: $(\Omega, \mathscr{F}, P)$
  - Ω spazio campione, i possibili esiti di un esperimento
  - $\mathscr{F}$ spazio degli eventi, famiglia di sottoinsiemi di Ω
  - $P : \mathscr{F} \rightarrow [0, 1]$, funzione di probabilità
- Variabile casuale X
  - $X : \Omega \rightarrow \mathbb{R}$ tale che $\forall\, r \in \mathbb{R}$, $\{\omega \in \Omega : X(\omega) \leq r\} \in \mathscr{F}$
  - $F_X(x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) = P(X \leq x)$
- Funzioni di distribuzione e di densità
  - Caratterizzano una variabile casuale
  - $F_X(x)$ è sempre valida nella sua definizione di base
  - Altre formulazioni, dipendono dal tipo della vc, discreta o continua

Distribuzioni discrete

- Variabile casuale discreta $X: \Omega \rightarrow x_1, x_2, \ldots, x_k, \ldots$
- Funzioni di densità e di distribuzione discrete
  - $f_X(x) = P(X = x)$ se $x = x_i$ per $i = 1, 2, \ldots k, \ldots$
    $f_X(x) = 0$ altrimenti
  - $F_X(x) = P(X \leq x) = \sum_{x_i \leq x} f_X(x_i)$
- Media, varianza e deviazione standard
  - $\mu_X = E[X] = \sum_i x_i f_X(x_i)$
  - $\sigma^2_X = Var[X] = \sum_i (x_i - \mu_X)^2 f_X(x_i)$

Variabili casuali e generatori

- Una variabile casuale è una mappa
  - Rappresenta un fenomeno non prevedibile
  - Associazione di un fenomeno a una variabile casuale
  - Analisi del fenomeno (della variabile), tramite osservazioni
  - Riproduzione della var. (del fenomeno) in una simulazione
- Distribuzioni uniformi discrete $X: \Omega \rightarrow 1, 2, \ldots, n$
  - $f_X(x) = P(X = x) = 1/n$ se $x = 1, 2, \ldots, n$
    $f_X(x) = 0$ altrimenti
  - Sappiamo come costruirne generatori (pseudocasuali)
- Distribuzioni continue
  - Definirle in generale, presentarne alcune notevoli
  - Definire metodi di generazione, in generale e su casi notevoli

Distribuzioni continue

- Variabile casuale continua: $X: \Omega \rightarrow \mathbb{R}$
- Funzioni di densità e di distribuzione continue

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du \qquad f_X(x) = \frac{dF_X(x)}{dx}$$

- Media e varianza

$$E[X] = \mu_X = \int_{-\infty}^{+\infty} x\, f_X(x)\,dx$$

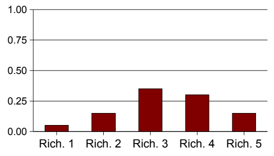
$$Var[X] = E[(X - \mu_X)^2] = E[X^2] - (E[X])^2$$

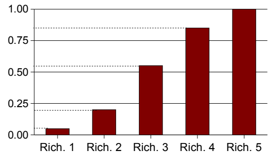
Generazione di dist. non uniformi

- Utilizzando distribuzioni uniformi
- Metodo del cappello (per distribuzioni discrete)
  - Procurarsi un buon numero di gettoni marcati con gli x_i
  - Metterne nel cappello in proporzione a $f(x_i)$
  - Estrarre con reinserimento
- Gli esiti riproducono la distribuzione (campionaria)
- In pratica
  - Generare un valore pseudocasuale
  - Riportarlo sull'asse verticale della funzione di distribuzione
  - Determinare il valore di x_i corrispondente

Esempio: tipi di richieste

- Percentuale delle richieste all'impiegato (da un campione)
  - Richiesta di tipo 1 05%
  - Richiesta di tipo 2 15%
  - Richiesta di tipo 3 35%
  - Richiesta di tipo 4 30%
  - Richiesta di tipo 5 15%





Metodo dell'inversione

- Generalizzazione del metodo del cappello
  - Funzione di distribuzione nota
  - E invertibile
- Inversione della funzione di distribuzione

$$U=F_X(x)=\int_{-\infty}^{x} f_X(u)\,du$$

$$x=G(u)=F_X^{-1}(u)$$

- In pratica
  - Generare un valore pseudocasuale
  - Applicare l'inversa e prendere il risultato

Metodo degli scarti

- Scartare i valori che non rispettano la distribuzione
  - Se la funzione di densità è nota e limitata in un dominio
  - Deriva dal metodo Monte-Carlo di integrazione
- In pratica
  - Data $f(x)$ definita in $[a, b]$
  - Definire c per riportare $f(x)$ in $[0, 1]$
  - Generare due valori pseudocasuali u_1 e u_2
  - $x = a+u_1(b-a)$ è accettato se $u_2 \leq cf(x)$
- L'efficienza dipende da quanto è "inscatolata" $f(x)$

Distribuzione esponenziale

- Distribuzione continua nei reali positivi

$$f_X(x,\lambda)=\lambda e^{-\lambda x} \qquad F_X(x,\lambda)=1-e^{-\lambda x}$$

- Media e varianza

$$E[X]=\frac{1}{\lambda} \qquad Var[X]=\frac{1}{\lambda^2}$$

- Caratteristiche
  - Si presta a modellare il tempo fra eventi successivi
  - Legata alla distr. discreta di Poisson (n. eventi nel tempo)

Generatore di dist. esponenziale

- Inversa della funzione di distribuzione

$$F_X(x,\lambda)=1-e^{-\lambda x}$$
$$u=1-e^{-\lambda x}$$
$$1-u=e^{-\lambda x}$$
$$\ln(1-u)=-\lambda x$$
$$x=-\frac{1}{\lambda}\ln(1-u)$$

- Note
  - $1-u$ e u sono equivalenti quando u è uniforme in $[0, 1]$
  - $1/\lambda$ è la media

Distribuzione normale

- Distribuzione continua nei reali

$$f_X(x,\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{(x-\mu)^2}{2\sigma^2}}$$

- Caratteristiche
  - Si presta a modellare gli scarti rispetto a un obiettivo
  - È definita la funzione di densità, μ è la media e σ^2 la varianza
  - Con μ = 0 e σ^2 = σ = 1, è la dist. normale standard N(0, 1)
  - Data z che sia N(0, 1), x è N(μ, σ) se

$$x=\frac{z-\mu}{\sigma}$$

Generatore di dist. normale

- Problema
  - Non è data la $F_X(x)$, il metodo dell'inversa non aiuta
  - La $f_X(x)$ è definita su tutto ℝ, il metodo degli scarti non aiuta
- Box-Müller: da cerchi uniformi a cerchi normali
- In pratica, per generare N(0, 1)
  - Generare due valori pseudocasuali u_1 e u_2 in (-1, 1)
  - Scartare le coppie fuori dal cerchio unitario, scalare il raggio

$$R=u_1^2+u_2^2<1 \qquad r=\sqrt{\frac{-2\ln(R)}{R}} \qquad x_1=u_1 r,\ \ x_2=u_2 r$$

  - I valori x_1 e x_2 sono N(0, 1)

Distribuzione Weibull

- Una distribuzione malleabile
  - Parametro di forma $k > 0$
    - $k = 1$, distribuzione esponenziale
    - $k = 2$, distribuzione di Rayleigh
    - $k = 3.4$, simile alla distribuzione normale
  - Parametro di scala $\lambda > 0$
- Funzione di densità (per $x \geq 0$) e di distribuzione

$$f_X(x,k,\lambda) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} \qquad F_X(x,k,\lambda) = 1-e^{-(x/\lambda)^k}$$

- La funzione di distribuzione è invertibile

Riferimenti

- M. Pidd, *Computer Simulation in Management Science*, Capp. 11, J. Wiley & Sons, 1998
- G. Gallo, *Note di Simulazione*, capp. 3.1 e 4.4
- Numerical Recipes in C, Cambridge University Press http://www.nrbook.com/